Information Systems and Economics (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/infse.2025.060103
ISSN 2523-6407 Vol. 6 Num. 1

Business and Finance Data Engine: A Solution and Practical Implementation for Financial Digital Transformation

Xu Rong

Shanghai National Accounting Institute, Shanghai, China

Keywords: Big Data, Business and Finance Integration, Configuration, Financial Digitalization

Abstract: The conversion of business and finance data is a critical and challenging aspect of financial digital transformation. This article conducts research based on the difficulties encountered in the digitalization of corporate finance, such as big data, development timeliness, and real-time data processing. Drawing on years of experience in financial digitalization, the article presents a solution that has been implemented within the author's company. It introduces the main functions of the enterprise's finance data engine, which aims to address the problem of converting business and finance data during the digital transformation of enterprises. The platform achieves data source access, data conversion rules, task execution, and data push through configuration, enhancing the flexibility and accuracy of data processing. It ensures that business data is generated in a timely, accurate, and complete manner to support the construction of a financial digital system.

1. Introduction

1.1 Background Introduction

In recent years, I research on the business of internet enterprises and how they achieve business and finance integration has revealed several key characteristics:

①Large data volume:Compared to traditional enterprises, these companies have a large number of orders with small individual transaction amounts, generating millions or even billions of data records daily.

②Strong business innovation capability:Internet companies iterate their versions every Tuesday or Thursday, or even more frequently. Behind this is a strong business team that can quickly respond to new market scenarios. Even in the absence of new market scenarios, they will iterate based on user or internal team feedback.

③Difficulty in fully grasping business details by finance:The number of finance personnel is much smaller than that of business personnel, making it difficult for finance to grasp every detail of business operations, even when business requirements are synchronized with finance. As a result, finance cannot ensure that accounting is correct.

④ Accounting time requirements: Accounting must be completed at a fixed time each month,

but business iteration is ongoing, and the update of the finance system lags behind the business system. Almost, the business system design must be completed before the financial system can be modified according to the rules, which may result in new business, split business, and abandoned business not being recorded on time. In addition, in order to meet the business analysis requirements for financial data, financial personnel need to issue financial data from month-end to T+1, or even real-time, to enable the business to make decisions on whether to continue promoting the business.

⑤ Complex billing rules: Traditional financial calculation is small, and Excel can meet the calculation needs, or even a regular calculator can meet the accounting needs. However, for internet enterprises, their billing rules require a lot of data to support calculations, such as the full settlement loan balance, which requires calculating the loan balance from the first loan.

⑥ Large volume of vouchers: When transferred to financial software, it is necessary to merge vouchers according to the consolidation rules. The traditional business and finance reconciliation, bank balance adjustment table, and account aging management logic are all unsuitable. It is necessary to redevelop the system to meet the needs.

Based on the new situations encountered, new problems need to be solved. Starting six years ago, I have been practicing a new business and finance integration system that can solve the above problems. After several years of practice and continuous iteration of new situations, I have finally completed a relatively complete system this year, which I call the Business and Finance Data Engine.

1.2 Research Purpose and Significance

This study aims to explore how large-scale transaction enterprises can better achieve the conversion of business data into financial data under the background of large-scale transactions and big data. Through the research on the construction, key technologies and applications, implementation strategies, and effectiveness evaluation of the Business and Finance Data Engine, this study provides enterprises with an efficient, scalable, and flexible[1] method for handling and analyzing business and finance data, thereby improving financial automation capabilities.

2. Construction of the Business and Finance Data Engine

2.1 Objectives of the Business and Finance Data Engine

According to the overall planning of the financial system, the Business and Finance Data Engine is the core of the enterprise's financial system. Its positioning is to connect business systems, acquire business data, convert business data into financial data, and provide the converted financial data to other systems for use. Upon completion of the entire financial system construction, it will form a platform (Business and Finance Data Engine) with four system configurations (Accounting, Management Reporting, Settlement, and Funds). For details, refer to Figure 1 in the Financial System Construction Plan.

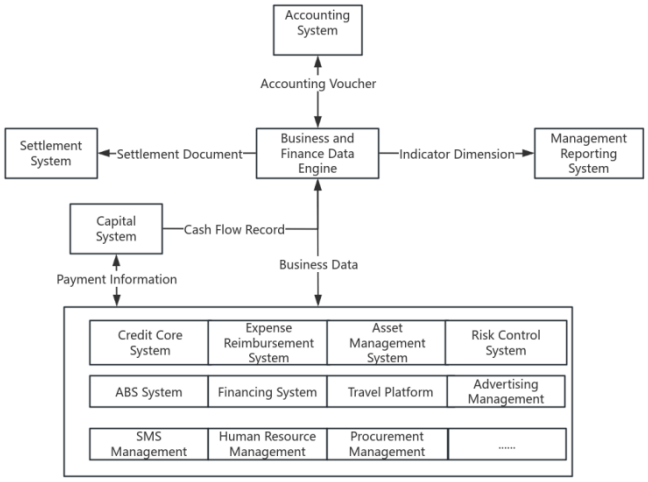

Figure 1: Financial System Construction Plan

Based on the above positioning, the objectives of the construction of the Business and Finance Data Engine are threefold:

(1) Build a Business and Finance Data Platform to enhance financial automation.

Firstly, the system should be capable of quickly integrating business data and processing it by cleaning, deduplicating, comparing, and storing the received data. Secondly, it should encapsulate corresponding financial scenarios based on business and finance scenarios, which include accounting needs, settlement requirements, and financial analysis requirements.

Again, the encapsulated business and finance scenarios should be set up to run batch tasks according to the timing requirements for different scenarios. After batch processing, the system should be able to query financial data, which can be traced back to the business data and allow for operations on the results.

Finally, the generated financial data should be regularly pushed according to the requirements of downstream systems. This enables other systems to further process the financial data, such as calling the clearing system in a timely manner for generated clearing data. The downstream results data are then fed back to the business system to facilitate the next steps in business processing, thus completing the nested financial processes within the entire enterprise process and achieving an enterprise business closure.

(2) Enhance the accuracy and self-consistency of financial data to prevent the reporting of incorrect data.

Data accuracy and self-consistency are core requirements for financial data. Different norms of accounting data, settlement data, and analysis data should be reconciled through internal correlation relationships to ensure the self-consistency of financial data, preventing the reporting of discrepancies in financial data that could lead to indecision or incorrect decisions. The results of the reconciliation should be automatically alerted through messages, prompting operators to take appropriate actions to promptly identify data issues and prevent the reporting of incorrect data.

(3) Reconcile business and finance data to provide data support for business development.

After achieving the above two objectives, business data and financial data can be reconciled. This avoids multiple abstractions in the process of converting business data to financial data, and the loss of multi-dimensional information contained in business data. It enables a financial perspective on business conditions, helping to identify low-value activities in business. From a business perspective, a more comprehensive view of the company's value can be gained. This, in turn, provides managers with more comprehensive information to assist in making scientific

decisions.

2.2 Business and Finance Data Engine Architecture and Data Flow Process

As mentioned earlier, the construction objectives of the enterprise's Business and Finance Data Engine and its core position within the financial system have been clarified. Based on this, the architecture of the Business and Finance Data Engine is depicted in Figure 2. The core modules of the Business and Finance Data Engine include: Master Data Management, Source Data Management, Target Data Management, Conversion Engine Management, Backend Task Management and Reconciliation Management, Data Push Management, and Document Management.

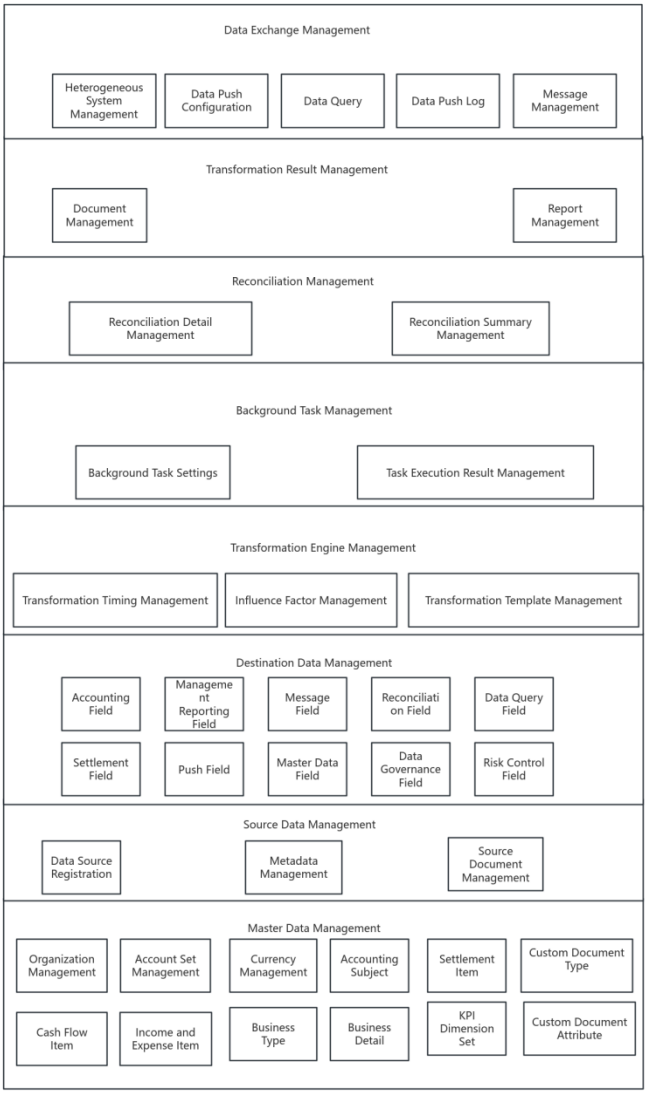

Figure 2: Business and Finance Data Engine Architecture Diagram

2.2.1 Master Data Management

Master data is the prerequisite for generating financial data. It is essential to ensure the consistency and accuracy of master data. Otherwise, it is impossible to generate financial data, or incorrect financial data may be generated. According to the needs, the system categorizes master

data management into two types:

The first type includes master data generated by other systems that the finance department needs to use, such as customer information, supplier information, material information, payment status, document status, etc., which describe real-world objects and enumerated data defined by the system. These data are beyond the control of the finance department, with a wide variety that cannot be exhaustive. The Business and Finance Data Engine uses custom file types to define data categories and metadata describing the data of each category. It manages the specific content of each data category through custom file attributes. This type of data requires the use of a conversion engine for data reconciliation, cleaning, conversion, and consolidation to ensure consistency and accuracy.

The second type involves master data managed by finance personnel, such as accounting subjects, customers, suppliers, etc. This data requires the creation of dedicated pages for specific management.

2.2.2 Source Data Management

Source Data Management forms the foundation of the Business and Finance Data Engine. It is responsible for collecting data from various data sources. Data collection and integration need to support multiple data formats and protocols, such as XML, JSON, CSV, etc., as well as various databases and file systems like MySQL, Oracle, Tidb, Hive. It supports various methods for receiving external system data, including FTP, messaging, interfaces, and binlog. Source data management mainly includes data source configuration, which sets up the source system, database, protocol, and interface of the data source to ensure connectivity with business systems. Metadata management configures which tables and fields to receive and pull from business systems.

2.2.3 Target Data Management

Target data refers to the type of financial data that business system data needs to generate, or intermediate data. The main function is to define the fields of target data, their types, whether they are referenced, whether they are displayed in lists and details, the display order, and whether they are query conditions, etc. The fields of target data can be adjusted as needed to ensure that finance can flexibly define the data structure according to requirements.

2.2.4 Conversion Engine

The Conversion Engine is the core of the Business and Finance Data Engine. It is responsible for processing, analyzing, and converting the collected data to generate valuable financial information. Its main functions include configuring the timing of financial data generation, encapsulating business scenarios, and combining business scenarios with corresponding financial data to form conversion templates. When converted into executable programs, the conversion templates support execution engines like HVIE, TIDB, MySQL, and JAVA based on data volume and business scenarios.

2.2.5 Reconciliation Management

Reconciliation is an essential part of financial activities, primarily to ensure data accuracy and consistency. The reconciliation management functions include:

Business-to-Business reconciliation, which is mainly used for the reconciliation of business data obtained by the Business and Finance Data Engine with business system data to ensure the correctness of source data.

Account-to-Real reconciliation, mainly used for the reconciliation of generated financial data

with business data to ensure the accuracy and completeness of the generated financial data.

Account-to-Account reconciliation, mainly used for the reconciliation of financial data generated by the Business and Finance Data Engine with financial data pushed to downstream systems to ensure the completeness and accuracy of the pushed data. The reconciliation results can also be queried.

2.2.6 Backend Task Management

Backend Task Management mainly configures scenarios that support real-time, scheduled data stream processing or distributed data batch processing. Its main functions include setting the frequency and execution time of conversion tasks and passing time parameters and field parameters to the conversion templates based on the configuration to filter the data range. After task execution, the results and log records are displayed. Management of result data can also be performed, such as tracing and querying business data, modifying data, cancelling task records, and rerunning batches, etc.

2.2.7 Conversion Result Management

Conversion Result Management serves as the user interface for displaying and applying financial data. It is responsible for presenting processed data to users in the form of charts, reports, etc. It is used for querying generated financial data and the corresponding business data. The generated data can be managed, such as invalidating, adding, and sending reports.

2.2.8 Data Exchange Management

Data Exchange is primarily used for pushing data generated by the Business and Finance Data Engine to other systems. It is achieved through configuring the data exchange methods between heterogeneous systems and the Business and Finance Data Engine (such as HTTP, MQ, FTP, etc.), as well as the data formats (such as JSON, TXT, etc.). This enables the implementation of data push functions, including push logs, push retry mechanisms, and data query capabilities, ensuring flexible configuration of data push and data consistency.

3. Application of the Business and Finance Data Engine

3.1 Application of the Business and Finance Data Engine in Business and Finance Integration

M Company's business encompasses various aspects, including proprietary loans, joint loans, and assisting loans. In addition to the core business operations that require accounting, there are also regular financing, financial bonds, ABS, membership, marketing, fund settlement, asset settlement, prepayment of service fees, commission prepayment, fund transfer, expense reimbursement, procurement, fixed assets, low-value consumables, leasehold interests in multiple workplaces, risk control data fees, SMS usage fees, server and related equipment leasing, and traffic usage fees, as well as staff salaries and performance bonuses. The daily production data volume for these businesses exceeds tens of millions of records, with complex interrelationships and a wide range of scenarios. To address the demands of the finance team and expedite the reduction of labor intensity, M Company decided to prioritize the automation of loan disbursement and repayment accounting. Subsequently, other business processes would be automated as conditions matured. The following section illustrates the application of the Business and Finance Data Engine in the accounting process using the loan disbursement and repayment business as an example.

3.1.1 Sorting out business scenarios

The method of sorting out business scenarios generally adopts the trace method. The trace method involves recording all business scenarios involving finance at each step of the process, moving forward from the beginning. This is usually done using mind maps to organize the results. M Company's proprietary lending and assistance lending businesses are relatively simple compared to syndicated lending, which encompasses both proprietary and assistance lending scenarios. Therefore, this article uses syndicated lending as an example to explain the process of sorting out business scenarios.

(1) Accounting Part: When accounting, as the accounting subject is M Company, only M Company's scenarios are sorted out. Sorting out the entire business process, ① Lending, when M Company transfers the right to use funds, cash flows from M Company to the customer and needs to be recorded. ② Repayment, which is divided into normal repayment and abnormal repayment. Normal repayment is further divided into actual repayment and prepayment. Regardless of actual repayment or prepayment, service fees will be charged to the joint capital providers according to the rules. Abnormal repayment is divided into early repayment and prepayment. Early repayment, in addition to principal and interest, also involves early handling fees. After the second day of overdue repayment, a report is filed, and the guarantor company makes a claim, transferring the creditor's rights from the capital provider to the guarantor. There may be recourse if repayment occurs, requiring the transfer of funds. There may also be scenarios of relief. Repayment scenarios, in addition to the above, may also involve member benefits, interest discounts, marketing activities, and other scenarios. ③ ABS debt, ABS is a way for our company to use debt for financing. When investors purchase, the debt needs to be recorded. This is a summary of the entire business scenario.

(2) Settlement Part: Settlement rule sorting is mainly done by settlement operations, with technical staff assisting. It is sorted out in combination with contracts and business processes. According to the business type: syndicated lending, assistance lending. Settlement objects: capital providers, guarantors, service providers. Billing objects: user repayments, risk data usage, advertising, SMS volume, etc. The results include settlement objects, billing scenarios, billing rules, and settlement periods.

(3) Management Reporting Part: M Company's management reports cover the entire company process, mainly divided into pre-event analysis, ①Contract signing decision-making based on the terms of the pre-signed contract to predict cash flow, bad rate, early repayment rate, and contract gross margin. Contracts are confirmed based on similar contract information. ② Loan volume control, based on cash flow forecasts, controls the credit limits of the company's and capital providers' loan accounts. According to the company's gross margin and bad rate control rules, controls the type of product a certain type of customer uses to meet, to ensure a balance between bad rate and gross margin. ③ Financing decision-making, analyzes the ABS cash flow to predict the amount of circulation in the circulation period to ensure normal payment in the payment period. Mid-event analysis, ① Analyzes the company's business performance, predicts the company's cash flow based on the company's lending and repayment plan. Analyzes the bad rate, early repayment rate, recovery rate of a product based on actual customer repayment. ② Provision calculation, provisions are made for service fee income based on the product's bad rate and early repayment rate, prudently confirming income. Post-event analysis, ① Pre-actual analysis, compares pre-event and mid-event analysis to find data gaps, improving the analysis model. ② Multidimensional business and finance reports, according to the company's assessment and management needs, issue multidimensional reports by product line, etc. Due to the large number of reports and data sources, the previous method of using SQL queries to collect data and then processing data step by step in

Excel had problems of low efficiency, slow report generation, and inaccurate data.

3.1.2 Financial Data Corresponding to Business Scenarios

Accounting Part: After sorting out the business scenarios, financial personnel and technical staff work together to match the financial data corresponding to the business scenarios. The elements of financial data generation include: ① Confirmation timing ② Measurement rules ③ Recording fields. The batch processing time for each scenario is confirmed according to the financial data requirements.

Settlement Part: The sorting of settlement rules is mainly done by settlement operations, with technical staff assisting. It is sorted out in combination with contracts and business processes. According to the business type: syndicated lending, assistance lending. Settlement objects: capital providers, guarantors, service providers. Billing objects: user repayments, risk data usage, advertising, SMS volume, etc. The results include settlement objects, billing scenarios, billing rules, and settlement periods.

Management Reporting Part: According to the analysis theme, sort out the analysis scenarios and the corresponding indicator dimension sets required by the scenarios. Sorting specifications: ① The indicator dimension set should be set according to the finest dimension. For example, the cost of repayment, product dimensions, etc., should be sorted out at the most detailed level to facilitate downstream BI systems for processing and display. ② Standardize the coding and naming of indicator dimensions, which should be easy to understand and avoid ambiguity. ③ Clarify categories, with business and finance layers divided according to business layers such as loan and repayment categories, daily management categories, etc., and finance layers such as income, cost, fund payment and receipt, provision, assets, liabilities, etc. After completing the indicator dimension set, summarize and organize it.

3.1.3 Sorting Out the Databases and Tables Required for Business Scenarios

Confirm the source databases, tables, fields, and values of the required fields based on the business and finance sorting results, and confirm the data acquisition methods[2]. Data synchronization frequency, etc. Ensure comprehensiveness and reliability during the sorting process. Comprehensiveness: The business data obtained can cover all business scenarios and provide data support for the encapsulation of settlement rules and settlement documents. Reliability: The data acquisition method can meet the requirements of timeliness and accuracy[3], and there is a corresponding compensation mechanism in case of abnormalities, providing a prerequisite for data conversion.

3.1.4 Business and Finance Data Engine Configuration

Configure the source system, metadata, and source documents according to the required databases, tables, and fields. Configure the conversion timing, influencing factors, and conversion templates based on the business scenarios and financial data comparison tables and financial data generation elements. Configure background tasks and run batches according to the financial data requirements period. Configure push rules according to the downstream requirements for the timeliness of financial data, timely meet downstream monthly settlement, fund payment and receipt, issuance of management reports, etc.

3.1.5 Data Comparison Verification

Offline manual financial data and online automated financial production data need to run in

parallel for three months according to the requirements of computerized accounting, and through comparison and verification, logical vulnerabilities in online logic are discovered to prevent production data errors after switching.

4. Conclusion and Prospects

4.1 Research Conclusion

This paper, set against the backdrop of high transaction volumes and big data, summarizes the difficulties and pain points in the conversion of business data to financial data. In order to address this issue, a Business and Finance Data Engine was designed, developed, and applied. Its full implementation at M Company has confirmed that the Business and Finance Data Engine can achieve the integration of business and finance under various scenarios, large transaction volumes, and big data. By comparing the efficiency, work intensity[4], and accuracy of financial data conversion before and after the system went live at M Company, the following conclusions can be drawn: The Business and Finance Data Engine can be configured to achieve multi-data source integration, data cleaning, business scenario encapsulation, business and finance data conversion, data reconciliation, and automatic financial data push based on demand periods. The process of system construction not only fulfills the financial data processing requirements but also explores a generalized business and finance data conversion solution, which can provide reference and inspiration for the integration of other enterprises' business and finance.

4.2 Future Research Directions

Looking forward to the future, the development trends of the Business and Finance Data Engine are as follows:

(1) Technological innovation and integration: With the continuous advancement of big data and artificial intelligence technologies, the Business and Finance Data Engine requires more technological innovation and application practice, which will further enhance the performance and functionality of the Business and Finance Data Engine.

(2) Deepening of industry-specific solutions: Enterprises in different industries will deepen the industry-specific solutions of the Business and Finance Data Engine according to their own business characteristics and needs, improving the efficiency and quality of converting business data to financial data. In particular, more effort needs to be invested in studying how to discover value from processed data for enterprises.

(3) Strengthening empirical research: Through large-scale empirical research, the application and effects of the Business and Finance Data Engine will be verified, better serving enterprises and improving their data management and decision support capabilities.

References

[1] Jin Yuan, Quanna Jun. Financial Data Middle Office Based on RPA Technology: Framework, Practice, and Reflection [J]. Finance and Accounting, 2024. (22): 17~23.

[2] Jin Yuan, Li Chengzhi. Research on Data-driven Financial Decision Support [J]. Accounting Communication, 2023. (3): 140-147.

[3] Li Duanmin, Chen Qiangyu. Research and Practice on Comprehensive Construction of Financial Data Based on Digital Transformation [J]. China Chief Accountant, 2021 (9): 138~140.

[4] Zhang Qinglong. Data Middle Office: Making Financial Data Usable [J]. Finance and Accounting, 2022 (9): 15~19.